

(19)



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: **07175495 A**

JPA 07-175495

(43) Date of publication of application: **14.07.95**

(51) Int. Cl.

G10L 3/00
G10L 3/00
H04M 3/42

(21) Application number: **05322662**

(22) Date of filing: **21.12.93**

(71) Applicant: **NIPPON TELEGR & TELEPH
CORP <NTT>**

(72) Inventor: **YADA KENICHI
YAMAMOTO HISAO
NAGASHIMA HIROMI**

(54) VOICE RECOGNITION SYSTEM

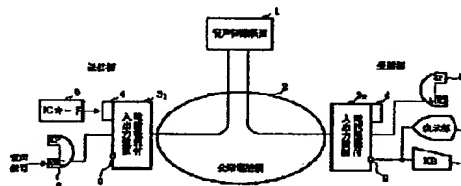
(57) Abstract:

PURPOSE: To utilize a voice recognition device between plural users through a communication network by respectively possessing memories storing own voice characteristic parameters for plural users, connecting the memories to communication terminals attachably/detachably and transferring the voice characteristic parameter stored in the memory to a voice recognition device when the voice recognition device is utilized.

CONSTITUTION: An IC card 5 is connectable to an input/output device with telephone function 3₁ being the communication terminal of a public telephone network 2, and the voice recognition device 1 is connected to the public telephone network 2, and a sound signal is transmitted from the input/output device with telephone function 3₁ through the public telephone network 2. Further, in the voice recognition device 1, a sound element standard pattern learning part 59 is provided with a means generating the voice characteristic parameter according to the inputted sound signal and the means transferring the generated voice characteristic parameter to the input/output device with telephone function 3₁, and the means writing the voice

characteristic parameter in the IC card 5 is provided in the input/output device with telephone function 3₁.

COPYRIGHT: (C)1995,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-175495

(43) 公開日 平成7年(1995)7月14日

(51) Int. Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G10L 3/00	551	A		
	521	A		
H04M 3/42		R		

審査請求 未請求 請求項の数 3 O L (全5頁)

(21) 出願番号 特願平5-322662

(22) 出願日 平成5年(1993)12月21日

(71) 出願人 000004226

日本電信電話株式会社

東京都千代田区内幸町一丁目1番6号

(72) 発明者 矢田 健一

東京都千代田区内幸町一丁目1番6号 日

本電信電話株式会社内

(72) 発明者 山本 尚生

東京都千代田区内幸町一丁目1番6号 日

本電信電話株式会社内

(72) 発明者 長島 広海

東京都千代田区内幸町一丁目1番6号 日

本電信電話株式会社内

(74) 代理人 弁理士 井出 直孝 (外1名)

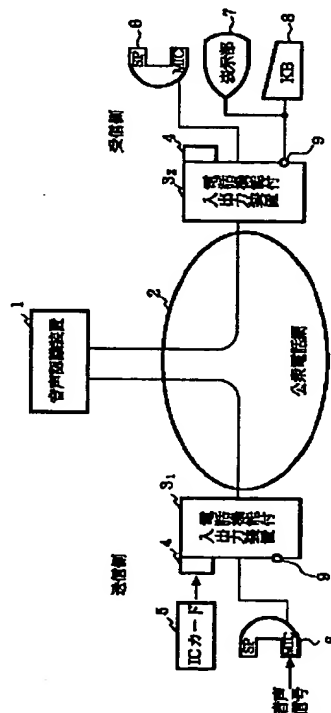
(54) 【発明の名称】 音声認識方式

(57) 【要約】

【目的】 通信網を介して複数の利用者が共同利用できる音声認識装置を実現する。

【構成】 音声信号は話者によりそれぞれ音声特徴パラメータが異なる。この音声特徴パラメータを各利用者がメモリに格納して所持し、この利用者が話者となって音声認識装置を利用したいときに、通信回線を介して音声認識装置にこの音声特徴パラメータを転送する。音声認識装置はこの転送された音声特徴パラメータにしたがって、音素標準パターンを作成し、音声信号を文字データに変換する。

【効果】 利用者が、それぞれ高性能な音声認識装置を個々に備えることなく、コストパフォーマンスの良い音声認識・文字データ変換システムを構築することができる。



できる。

【 0 0 1 2 】

【実施例】本発明実施例の構成を図 1 および図 2 を参照して説明する。図 1 は本発明実施例の全体構成図である。図 2 は本発明実施例装置のブロック構成図である。

【 0 0 1 3 】本発明は、音声特徴パラメータが格納された I C カード 5 と、この音声特徴パラメータを取込み入力された音声信号を文字データに変換する音声認識装置 1 とを備えた音声認識方式である。

【 0 0 1 4 】ここで、本発明の特徴とするところは、 I C カード 5 は通信網としての公衆電話網 2 の通信端末である電話機能付入出力装置 3_i に接続可能であり、音声認識装置 1 はこの公衆電話網 2 に接続され、音声信号は電話機能付入出力装置 3_i からこの公衆電話網 2 を介して伝送されるところにある。また、音声認識装置 1 には、入力された音声信号にしたがって音声特徴パラメータを生成する手段と、生成された音声特徴パラメータを電話機能付入出力装置 3_i に転送する手段とを音素標準パターン学習部 5 9 に備え、電話機能付入出力装置 3_i には I C カード 5 にその音声特徴パラメータを書込む手段を備えている。また、受信側の電話機能付入出力装置 3_j の入出力端子 9 には、表示部 7 およびキーボード 8 が接続されている。

【 0 0 1 5 】次に、本発明実施例の動作を説明する。図 1 に示すように、送信側の電話機能付入出力装置 3_i の送受器 6 から音声信号を入力して、受信側の電話機能付入出力装置 3_j に接続された表示部 7 に文字データを出力する場合について説明する。送信側は電話機能付入出力装置 3_i から音声認識装置 1 を呼び出す。送信側と音声認識装置 1 との呼設定が完了した後に、送信側は音声認識装置 1 からの指示にしたがい、文字データの送り先（受信側）のアドレス（電話番号）と送信側の話者の音声特徴パラメータを音声認識装置 1 に転送する。この音声特徴パラメータは I C カード 5 に格納され、接続端子 4 にこの I C カード 5 を接続することにより、電話機能付入出力装置 3_i から音声認識装置 1 への音声特徴パラメータの転送が可能になる。

【 0 0 1 6 】音声認識装置 1 は、電話機能付入出力装置 3_i から転送された音声特徴パラメータにしたがって音素標準パターンを作成し、送信側に音声信号入力を指示する。指示にしたがい、送信側の話者は音声信号を入力する。音声認識装置 1 は、音声認識を行い文字データへの変換を行う。このとき、音声認識装置 1 は認識結果を一旦送信側に提示して修正を行うようにすることもできる。文字データへの変換結果は、一旦蓄積した後に、電話機能付入出力装置 3_j を介して受信側に送信され、表示部 7 に表示される。前述したように、文字データは文字単体または文章として表現され、文字、数字、記号が含まれている。文字データは、受信側が情報の意味を認識できればいかなる形でもよく、例えば単なるアスキー

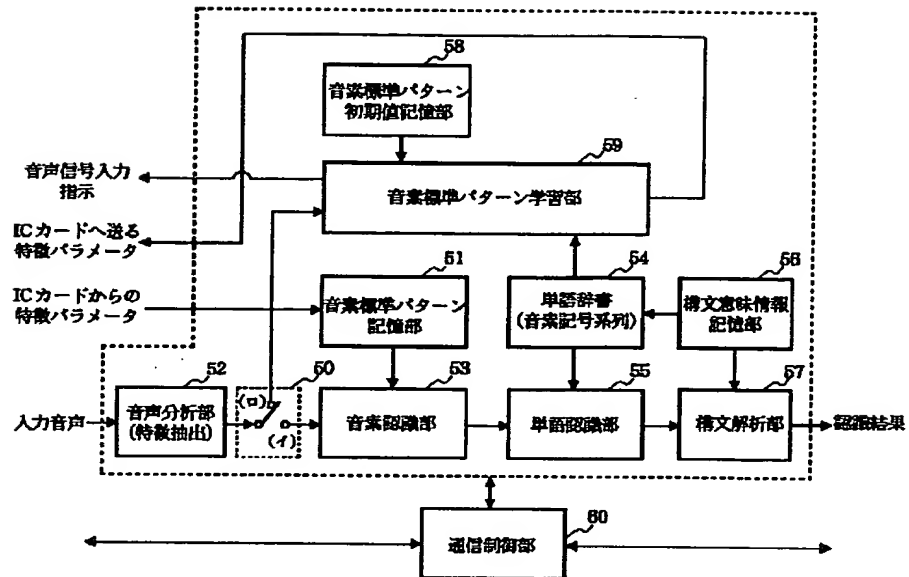
コード以外にも、通信端末毎に異なる暗号であってもよい。

【 0 0 1 7 】次に、音声認識装置 1 の動作を図 3 を参照して説明する。図 3 は通信制御部 6 0 の動作を示すフローチャートである。呼設定、その他の通信制御は、図 2 に示す通信制御部 6 0 により行われる。また、ここでは通信制御部 6 0 は、各ブロックの動作も制御しているとして説明する。音声特徴パラメータ作成過程と音声認識系とを切替えるスイッチ 5 0 は、通常は音声認識を行う端子（イ）側に接続されている（S 1）。電話機能付入出力装置 3_i からの音声特徴パラメータは音素標準パターン記憶部 5 1 に入力され（S 2）、パターンマッチング用のテンプレートとしての音素標準パターンが作成される（S 3）。このとき、入力される音声特徴パラメータは日本語すべてでも数 1 0 k b 程度でよく、I C カード 5 の容量は十分に音声特徴パラメータの格納に対応できる。また、転送時間も実用上問題にはならない。続いて、電話機能付入出力装置 3_i の送受器 6 のマイクロホンから入力された音声信号は、音声分析部 5 2 で音声特徴パラメータの時系列に変換され（S 4）、スイッチ 5 0 を介して音素認識部 5 3 に入力される。音素認識部 5 3 では、音素標準パターン記憶部 5 1 からの音素標準パターンと入力の音声特徴パラメータとの照合を行い、音素信号系列として出力される（S 5）。音素認識された音素記号系列は単語認識部 5 5 に入力され、単語辞書 5 4 を参照することにより語の切れ目毎に単語認識される（S 6）。さらに、単語認識結果は単語毎の音素記号系列として構文解析部 5 7 に入力され、構文意味情報記憶部 5 6 の構文および意味などに関する言語的知識に基づいて文章としての修正が加えられ、最終的な認識結果が出力される（S 7）。

【 0 0 1 8 】I C カード 5 に未だ音声特徴パラメータが格納されていない場合に、音声認識装置 1 は電話機能付入出力装置 3_i からの音声信号により音声特徴パラメータを作成し、これを電話機能付入出力装置 3_i の I C カード 5 に転送する。この場合には、スイッチ 5 0 を端子（ロ）側に接続して行う（S 1）。音素標準パターン学習部 5 9 からの、例えば「五十音を最初からゆっくりとしゃべってください。」などの音声信号入力指示により（S 8）、適当に決められた順序で送受器 6 から入力された音声信号は、音声分析部 5 2 で音声特徴パラメータの時系列に変換され（S 9）、さらに音素標準パターン学習部 5 9 で音素標準パターン初期値記憶部 5 8 からの初期値の音素標準パターンと単語辞書 5 4 により、音素境界のセグメンテーションとそれに基づく音声特徴パラメータの平均化が行われる（S 1 0）。結果は、再び電話機能付入出力装置 3_i を介して I C カード 5 に転送され（S 1 1）、音声特徴パラメータとして保持される。

【 0 0 1 9 】本発明実施例では、通信端末は I C カード 5 を着脱自在に収容できる電話機能付入出力装置 3_i 、

【図 2】



【図 3】

